DOCUMENT RESUME

ED 222 271 PS 013 128

AUTHOR Martin, Roy P.
TITLE Temperament: A Review of Research with Implications for Child Psychology in the School and Clinic.
PUB DATE Aug 82
NOTE 30p.; Paper presented at the Annual Meeting of the American Psychological Association (90th, Washington, DC, August 23-27, 1982).

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; Child Psychology; *Clinical Psychology; Cognitive Ability; Definitions; Elementary Education; *Genetics; Literature Reviews; *Personality; *Personality Measures; Personality Studies; Psychopathology; Student Adjustment

ABSTRACT

Briefly considering the literature, this paper reviews research in the area of temperament and points out the implications of findings for school psychologists. Following a brief discussion of several definitions of temperament, two issues given the most consideration in the theoretical literature are outlined: the genetic origins of temperament and the stability of temperament over time. A brief summary of empirical evidence bearing on both issues is presented. Next, research is reviewed which is devoted to the development of measurement devices for assessment of temperamental characteristics in infants, toddlers, preschool and elementary children, middle-school children, and adults. Finally, a selected review of current research examining the effects of temperamental differences on such areas as the development of psychopathology, school adjustment, intelligence, and academic achievement are summarized and discussed. (MP)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ED222271

Temperament: A Review of Research With Implications for Child Psychology in the School and Clinic

Roy P. Martin

University of Georgia

PS013128

Running Head: Temperament

Portions of this paper were presented at the annual meeting of the American Psychological Association, Washington, 1982.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Roy P. Martin
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Temperament: A Review of Research With Implications for Child Psychology in the School and Clinic

There has recently been a dramatic increase in research activity directed at determining the developmental, clinical, and educational correlates of temperament. This increased activity seems to be attributable to two general trends in current psychological thought. First, there is at present a strong impulse throughout psychology to look at biological determinants of behavior. The interest in clinic neuropsychology is one manifestation of this trend (e.g., Hynd & Obrzut, 1981), as is the increased activity in behavior genetics (e.g., Thompson & Fuller, 1978), and interest in biochemical correlates of developmental phenomenon (e.g., Maccoby, Doering, Jacklin, & Kraemer, 1979). Second, there appears to be some dissatisfaction with the explanatory power of the tabula rosa, exclusively environmental, models of individual differences. Parents who consistently report that their children were very different from birth, and teachers who report large individual differences in reaction to instructional programming and behavior management procedures, that are apparently unrelated to cognitive variables, are implicitly asking researchers to look for sources of individual differences other than those most frequently considered by psychologists. Psychologists have traditionally paid lip service to constitutional differences but with few exceptions have not directed their energies toward documenting such differences or examining the correlates and consequences of such differences.

There are currently a sizeable number of behavior geneticists looking at non-cognitive individual differences, frequently labeled temperamental differences, to determine the contribution of heredity to these differences. There are also several active research groups devoted to the development of measurement devices designed to assess temperamental differences. Finally, an even wider number of clinicians, developmentalists, and ecucational researchers are studying the effects of temperamental differences on such important areas as the development of psychopathology, school adjustment, intelligence and achievement.

It is the purpose of this paper to briefly review the highlights of this research, and to point out the implications of this research for school psychologists.

Toward A Definition of Temperament

Before reviewing temperament research, it is necessary for the reader to understand what is meant, and perhaps more importantly what is not meant, by the term as it is currently understood. Definitional clarity is particularly critical for the concept of temperament because the word has popular and historical connotations that are distinct from those considered by contemporary researchers. Unfortunately there is not one generally accepted definition of temperament. Thus, several definitions are presented for the purpose of acquainting the reader with the various emphasizes of writers on the subject.

Prior to current interest in temperament several attempts were made to differentiate temperament from personality. One such attempt was made by Allport (1937:

> "Temperament refers to the characteristic phenomena of an individual's emotional nature, including his susceptibility to emotional stimulation, his customary strength and speed of response, and the quality of his prevailing mood, and all pecularities of fluctuation and intensity in mood; these phenomena being regarded as dependent upon constitutional make-up, and therefore largely hereditary in nature (p. 54).

Cattell (1946), postulated three broad classes of individual differences phenomena of which one was temperament, and the other two were cognitive differences and motivational differences. One way in which Cattell's temperamental variables (excitability, sensitivity, perseveration, and impulsiveness) were differentiated from the other two classes of traits was that they were felt to be the least modifiable through environmental manipulation.

Current researchers and theoreticians utilize several aspects of these two definitions. Thomas, Chess, and colleagues, (Thomas & Chess, 1977; Thomas, Chess, Birch, Hertzig, & Korn, 1963) differentiate temperament from motivation by defining temperament as a behavioral style. In this context style refers to the "how" of behavior,

not the "what" or the "why." The "what of behavior is the content, such as the smile of an infant, or the cry of a preschooler. The "why" refers to the motivational component. The preschooler may cry because she has learned that to cry will bring a caretaker who will relieve discomfort. Style then refers to such variables as how intensely the child cries, what the threshold of discomfort is that produces crying, and the duration and latency of the crying. Thomas and Chess (1977) have researched nine temperamental variables which they feel qualify as stylistic variables: intensity, threshold, activity, rhythmicity, adaptability, approach/withdrawal, distractibility, persistence, and mood.

An alternate approach to defining temperament is presented by Buss and Plomin (1975), who established five criteria that they feel are required before a trait can be classified as temperamental: (a) there must be evidence the trait is inherited; (b) the trait must be shown to be somewhat temporally and situationally stable; (c) it must be predictive of adult personality; (d) the trait must be adaptive in an evolutionary sense; (e) the trait must be present in animals. Considering evidence on these points Buss and Plomin postulate that four broad traits can be classified as temperamental: emotionality, activity, sociability, and impulsivity.

Rothbart and Derryberry (1981) place central importance in their definition of temperament on the hypothesized link between central nervous system reactivity and behavior. In their theory temperament refers to individual differences in reactivity and self-regulation, both assumed to have a constitutional basis. Reactivity refers to the excitability, irritability and responsiveness of the neurological system of the organism. Self-regulation refers to neural and behavioral processes like attention and social approach/avoidance which the organism can use to increase or decrease the strength of environmental stimuli impinging on the central nervous system.

Distilled from the above we can see that temperamental variables are felt to be relatively stable traits, of genetic or congenital origin, that are descriptions of qualities of emotion, and characteristics of central nervous system arousal as they

are reflected in behavior. Activity and approach/withdrawal, for example, are temperamental traits that are assumed to be related to arousal, while intensity and threshold are characteristics of emotional responsiveness. In order to further clarify contempory positions on temperament it must be understood that current thought does not maintain that temperamental traits are immutable, or necessarily present at birth. Further, no direct connection between body type and temperament is postulated as earlier researchers did (Sheldon, 1942). Further, as measured at any given time in the development of the organism it is understood that the measured behavior is the result of a complex interactive function of genetic and environmental influences.

Two issues that have been given the most consideration in the theoretical literature on temperament are the issues of the genetic origins of temperament and the stability of temperament over time. Given the fundamental importance of these issues a brief summary of the empirical evidence bearing on both issues will be presented.

Temperament and Behavior Genetics

The most direct studies in behavior genetics, of course, take place in animals. In such studies animals are selected from an existing strain which seem to possess more of a given behavioral tendency (e.g., aggression) and other animals from the same strain are selected who possess less of the tendency. Through processes of repeatedly mating the animals possessing more of the tendency and also mating animals possessing less of the tendency two distinct substrains are produced. With some traits it has been found that after as few as ten generations, the two substrains are completely distinct. That is, if aggression were the trait under consideration, the least aggressive animal in the aggressive strain is more aggressive than the most aggressive animal in the non-aggressive strain.

Among the temperamental behaviors which have responded to such selective breeding are emotional reactivity, activity level, sexual competence (a form of sociability), and competitive aggression. These experiments have been carried out with species ranging from fruit flies to dogs (Thompson & Fuller, 1978). Diamond (1957) an animal researcher, postulated four basic temperaments that are shared by mammals close to man on the

phylogenetic continuum and man. These are aggressiveness, affiliation, fearfulness, and impulsivity. Diamond's theory and research has had an important impact on current thinking about temperament in humans (Buss & Plomin, 1975).

Behavior genetics research in man relies almost exclusively on twin studies in which correlations between monozygotic twins (twins who have identical heredity) and between dyzygotic twins (twins who do not have identical heredity but have very similar environments) on a personality or temperamental trait are compared. While this research rests on a variety of assumptions, and the mathematics of determining inheritability can be quite complicated and controversial, the basic notion is that if the correlation for monozygotic twins is significantly larger than those for dyzogotic twins, then the trait is assumed to be in part genetically determined. Using such techniques in their own research and reviewing a number of prior studies, Buss and Plomin (1975) come to the conclusion that there is strong empirical evidence for the inheritance of activity, sociability and emotionality in man. They found only weak evidence for the inheritance of impulsivity. In support of Buss and Plomin, Goldsmith and Gottesman (1980) have recently found evidence for a genetic component in activity level. Scarr (1969) focused her review on sociability and presented a convincing array of data to support the inheritance of this temperament trait. However, Thompson and Fuller (1978)in their review state that in the realm of temperament and personality unequivocal evidence for inheritability only exists for introversion-extraversion. Others feel that since introversion-extraversion is a complex factor made up of sociability and impulsivity, that these two traits are the more fundamental aspects of temperament. Thus, the evidence for the genetic component in introversion-extraversion is taken by many to demonstrate the genetic component in sociability and impulsivity.

To summarize, evidence exists for the inheritance of several broad temperamental traits in animals and in man. While different researchers give their variables somewhat different names, the traits for which there is the most clear evidence is activity, sociability and emotionality, with weaker evidence for impulsivity.

Temporal and Situational Stability

A "strong trait" theory of temperament would postulate that temperamental characteristics are present at birth and are consistently manifest throughout the life of the individual. Such a position is naive! It is contrary to a substantial amount of easily observed data, and it is unnecessary for a viable temperamental theory.

A "strong trait" theory would deny or minimize effects of the environment. However, it can be readily seen that even a relatively stable characteristic like height, which is known to be genetically determined, can be strongly influenced by diet and illness. Further, the expression of a temperamental trait is obviously more sensitive to environmental events than is height. Consider two classroom teachers, one who is relatively tolerant of activity in the classroom, and another in which the teacher emphasizes orderliness. Individual differences in the characteristic of activity level will be much more apparent in the first instance than in the latter. If the environmental press is strong enough, individual differences may be eliminated altogether.

The example given above focuses on situational stability, but a similar example could be given for temporal stability since time and situation are often interactive variables. A child may seem stable on the sociability dimension during the preschool years as rated by his parents because they see him in the home and other situations in which they are present. However, his sociability rating might change during the age period five to seven because at school without parental support the child adapts to social situations more slowly and initiates fewer interactions.

Another difficulty in obtaining temporal stability in temperament measures occurs because some genetically determined processes are not expressed in a consistent way from the time of birth. The timing and the specific manifestation of the complex mechanisms that bring about sexual maturity during adolescent are, in part, genetically determined but are not seen during childhood. It is apparent from this example that characteristics or traits may become apparent only at specified times during development. It is currently not know to what extent the onset and intensity of temperamental variables are controlled in this way.

Another problem in obtaining temporal stability of a behavioral trait is that with changes in age, the manner in which a temperamental trait is expressed can change radically. Activity level is manifest in infancy in a very different manner than it is in adolescence.

Perhaps the most appropriate way to conceptualize a temperamental trait is to think of it as a predisposition. Emotionality can be viewed in this way. Given specific environmental circumstances, emotionality may results in neurotic behavior. With time and change in circumstance, the behavior may dissipate. Persons observing the individual will be unable to see signs of the maladaptive fearful pattern of behavior that was once expressed, and may not be able to discern the more subtle signs of emotionality. Weeks, months, or years later, a stressful circumstance may produce the original neurotic behavior again. This example makes clear the problem of determining the stability of emotionality across time. The individual was perpetually in a predisposed state, but that state was not readily observed except in environments that precipitate a particular behavioral manifestation.

When all the factors mentioned above are considered it becomes clear that contemporary theorists subscribe to a "weak" or "interactional" temperament trait theory. Given such a theory strong empirical evidence of temperamental stability is not expected except over relatively brief periods of time or in the stable environments. In general the empirical results on the temporal and situational stability of temperament have been consistent with this expectation.

In the New York Longitudinal Study, the degree of consistency for each of the nine Thomas and Chess temperaments over the first five years of life were determined (Thomas & Chess, 1977). Statistically significant correlations were obtained between most categories for adjacent years. As the time span was increased to two and three years the number of significant correlations decreased. Greatest stability was seen for activity level, adaptability, and threshold.

Buss and Plomin (1975) reviewed empirical research on stability of the four temperament variables in their theory. For activity they found five studies which measured

activity longitudinally from infancy through childhood: All five reported the same finding; that is, little or no stability of activity across these ages. Two explanations are offered for this result. First, activity level during infancy is difficult to measure because infants are not often active. Second, there seems to be little comparability between active behavior in infancy and in childhood. For emotionality, most evidence shows weak to moderate stability during the first year. Research on older children shows moderate stability.

There appears to be little stability of sociability during the infancy period. However, responsiveness to persons and shyness show considerable stability over the toddler and early preschool years with these results being stronger for girls than for boys. Sociability from the later preshool period through adulthood seems quite stable with most coefficients being in the .40 to .70 range. This data led Buss and Plomin to conclude that sociability is the most stable temperament. Finally, there have been so few longitudinal studies that have measured impulsivity and related concepts (persistence, boredom, sensation seeking, decision time, inhibitory control) that is impossible to describe an empirical trend.

An interesting observation made by Eliasz (1980) is that persons with different temperamental patterns may be differentially stable across time and setting. Summarizing data from several studies Eliasz finds that highly reactive persons (persons with low thresholds for stimulation, for example, introverts) are less situationally stable in their behavior than less reactive persons (e.g., extroverts). Eliasz hypothesized that this occurs because more reactive persons are more sensitive to differences in social environments and the demands of these environments. More research of this type could be helpful in providing a better understanding of the meaning of stability in temperament research.

Measurement of Temperament

The Infant Period. There are three instruments that have been used fairly extensively to measure temperamental variables during the infancy period. The first one to be developed and the one most extensively used is the Infant Temperament Questionnaire

(Carey, 1970; Carey, 1972; Carey & McDevitt, 1978). In its current form it consists of 95 items measuring all nine of the Thomas and Chess temperament dimensions (activity, rhythmicity, adaptability, approach/withdrawal, mood, intensity, distractibility, persistence, threshold). Each item is rated on a six point continuum (1 = almost never, 6 = almost always). Items were originally developed from the interview procedure used by Thomas and Chess in the New York Longitudinal Study. Internal consistency reliabilities for the scales range from .49 to .71. Norms, factor structure, and cluster structure are available for infants seen by Carey in his pediatric practice in Pennsylvania. No manual summarizing reliability and validity data is currently available.

Rothbart (1981) has recently developed an infant scale designed to measure activity level, smiling and laughing behavior, fearfulness, distress to limitations, soothability, and duration of orientation. Items are responded to on a seven point Likert scale (1 = never, 7 = always) with an eighth choice being "does not apply." Rothbart describes the development of the scale and the results of item analyses and reliability studies on a large sample of 3 to 12 month olds. Internal consistency of the scales ranges from .49 to .69 and temporal stability (test-retest reliability) for duration ranging from three to nine months were generally in the .65 to .85 range. No normative data, factor or cluster analysis data, or validity data are reported.

A rating scale that is sometimes considered to include measures of temperament is the Neonatal Behavioral Assessment Scale (Brazelton, 1973; Als, Tronick, Lester & Brazelton, 1979). This scale measures neurological intactness by assessing the strength of 20 reflexes (e.g., rooting, Moro). Two global questions are also included which "summarize the neonates overall organization." These are the attractiveness and need for stimulation ratings. Finally, the examination assesses the newborns' behavioral repetoire on 26 items, grouped into four categories: interactive capacity; motoric capacities, organizational capacity in respect to state control, and organizational capacity regarding physiological responses to stress. The neonate is usually assessed twice, at two to three days after birth and again at nine to ten days. While some

temperament like data can be gleaned from the Brazelton procedure, the examination was not primarily designed to assess temperament and Goldsmith and Campos (1981) argue against its use for this purpose.

The Toddler Period. Fullard, McDevitt, and Carey (Note 1) have developed a temperament scale for the toddler period (ages 1 to 3). This 97 item instrument was designed to measure all nine Thomas and Chess temperament variables. The Toddler Temperament Questionnaire uses the same scoring procedure used in the other McDevitt-Carey devices. Like the other instruments in this series, the norms were developed from children in Carey's pediatric practice. Internal consistency coefficients for the scales range from .53 to .86 and test-retest coefficients range from .69 to .89. The questionnaire has not been extensively used to date, and the authors provide no validity data on the instrument. No manual is available.

The Preschool-Early Elementary School Period. In their 1977 book Thomas and Chess published the Parent and Teacher Temperament Questionnaires. Both are rating scales based on earlier structured interviews. The Parent Temperament Questionnaire was designed to measure activity, rhythmicity, adaptability, approach/withdrawal, threshold, intensity, mood, distractibility, and persistence. The scale contains 72 items, with eight items measuring each temperament trait. The Teacher Temperament Questionnaire consists of 64 items measuring all the traits measured by the parent scale with the exception of rhythmicity.

In an analysis of the item characteristics and reliability of the Parent Questionnaire, Martin (Note 2) found that the internal consistency of several scales was very low (four scales had alpha coefficients below .60). Second, there was a good deal of redundancy in the scales as indicated by correlations among scales.

Martin (Note 3) has also carried out an analysis of the Teacher Temperament Questionnaire on 401 children, the great bulk of whom, are in kindergarten and first grade. This analysis shows the teacher questionnaire to have relatively strong internal consistency characteristics and test-retest reliability. Several of the items scattered throughout

the questionnaire perform poorly, but the questionnaire in general is of sufficient reliability to support its use.

The Parent Questionnaire has been little used for research so the validity of the measure is uncertain. The Teacher Questionnaire has been used somewhat more frequently, but no summary of relevant research is now available.

In 1978, McDevitt and Carey presented initial data on a 108 item parent rating scale called the Behavioral Style Questionnaire (BSQ). This questionnaire was based on the Thomas and Chess interview procedure. Additional items were developed and assigned to temperamental categories based on interjudge agreement. The scale was normed and reliability studies were carried out on a sample obtained from Carey's pediatric practice. Reliability is generally adequate for a research instrument, however, internal consistency estimates for the rhythmicity scale, the threshold scale, and the persistence scale were .60 or below. The developers report cluster analysis data which supports the clusters found by Thomas and Chess (1977). That is, three clusters were found, the difficult child cluster, the easy child cluster, and the slow-to-warm-up child cluster. The BSQ has been extensively used in research on temperament of children age three through seven and a great deal of supportive validity data has been accumulated. Unfortunately, this data has not been summarized, and no manual for the instrument has been developed. Other weaknesses of the scale include a somewhat awkward scoring system, and norms based on a geographically and socio-economically restricted population.

An alternative to the Behavior Style Questionnaire and the Parent Temperament Questionnaire is the Parent Form of the Temperament Assessment Battery (Martin, Note 3). This scale is an adaptation of the Parent Temperament Questionnaire (Thomas & Chess, 1977). The scale was derived by eliminating items from the original instrument that had item-total correlations of .30 or below on each scale, and replacing them with items that were rationally derived to measure the construct that appeared to be measured by the remaining items. These items were further field tested and the process was repeated.

The instrument was designed from the beginning as an instrument for clinical or applied use, as well as research use. It was designed to overcome some of the limitations

of the BSQ and the Original Parent Temperament Questionnaire. In particular, efforts were made to shorten the scale as much as possible consistent with psychometric consideration. Second, several scoring systems were developed and tested in order to develop one which minimized scoring time and error. Third, subscales of questionable theoretical or psychometrical quality were eliminated.

In response to the last goal, the mood, rhythmicity, and threshold scales of the original scale were omitted. Thus, the Parent Form of the TAB contains activity, adaptability, approach/withdrawal, intensity, persistence, and distractibility scales. The internal consistency of these six scales range from .57 to .87. While these internal consistency figures are substantially stronger than the original Thomas and Chess scale and are comparable to those reported by McDevitt and Carey (1978) for the BSQ, they are still too low to allow the scales to be used in applied circumstances. However, exploratory factor analysis revealed a three factor solution with two scales loading highly on each factor and no scale loading highly on more than one factor. Based on this analysis three factor scales were derived and labeled emotionality, persistence and sociability. These three factors have internal consistency coefficients (alpha) of .77, .77, and .82, respectively. Since alpha is a conservative estimate of reliability, these scales are homogeneous enough to be of practical use in some situations. No test-retest reliability for the scale has yet been carried out, but several validity studies are reported in a manual which provides some evidence for the construct validity of the scale.

The Parent Form of the Temperament Assessment Battery (TAB) is one of three forms in the battery. The other two are the Teacher Form, and the Clinicians Form. The Teacher Form is a modification of the Teacher Temperament Questionnaire of Thomas and Chess (Thomas & Chess, 1977). This modification eliminates the mood and threshold scales, and several weak items on the original scale. Eliminated items were replaced by items that apparently measured a construct similar to the remaining items in the scale. These modifications are currently being field tested. Factor analysis utilizing a three factor solution produced factors that were similar to those produced for the Parent Form. Internal consistency of these factors was slightly higher than for the Parent Form, ranging

from .77 to .90. Test-retest reliability for the scales ranged from .53 for adaptability to .81 for persistence for a six month interval. Several validity studies have supported the construct validity of the scales.

The third questionnaire in the TAB is the Clinicians Form. This form was designed to be used by school psychologists, and other child psychologists who have observed a child in an assessment situation. The scale was designed to parallel the measurements obtained from the Parent Form and the Teacher Form. Validity and reliability data on this scale are currently being analyzed.

One major advantage of the TAB is that it allows for multi-source assessment (parent, teacher, clinician). This makes possible the control of the major weakness of ratings scales, that is, the subjectivity of the source of the rating. The battery also helps control for setting effects in that effects of the school environment, home and clinicians office can be observed in the ratings made of the child. The TAB is also accompanied by a manual which summarizes the psychometric properties of the instruments and validity studies carried out to date. The primary weakness of the TAB is that interrater reliability across forms (a comparison of the ratings of mothers and teachers, for example) tends to be low. It is not known at present whether this is due to a characteristic of the instruments, or the nature of the amount of variance introduced by different settings, and different raters. The interrater reliability problem tends to be a major problem for other rating scales, as well.

The Middle School Period. To augment the Infant Temperament Questionnaire, The Toddler Temperament Questionnaire, and the Behavioral Style Questionnaire, Hegvik working with McDevitt and Carey (Note 4) developed a temperament scale for middle childhood (ages 8 through 12 years). The scale consists of 99 items measuring eight of the nine Thomas and Chess dimensions (rhythmicity was omitted). In addition a scale measuring predictability or behavior consistency is also included. Some normative and reliability data are available with the instrument, but again the data comes only from subjects in Carey's pediatric practice. There has been little research using the instrument to date, so validity cannot be currently evaluated.

Adult Temperament Measures. In the adult period it is difficult to determine if a personality assessment device measures temperament or not because many omnibus instruments do not differentiate between temperament and personality scales. One measure which explicitly measures adult temperament is the EASI-III Temperament Survey of Buss and Plomin (1975). This scale is a self-report measure of emotionality, activity, sociability, and impulsivity. It is made up of 50 items that ask for broad direct judgments. A typical item in the impulsivity section is: "I have trouble controlling my impulses." Details on the scale are scattered through Buss and Plomin's book. The scale is reproduced in the appendix of the volume. A good deal of validity data is also reported. Other measures of some temperamental variables for adults are included in the Thurstone Temperament Survery, 16 Personality Factor Scale, Eysenck Personality Inventory, Guilford-Zimmerman Temperament Survey, The California Personality Inventory, and The Sensation Seeking Scale (Zuckerman, 1979).

Other Measures. All the measures discussed to this point were designed to assess temperamental characteristics in one developmental age group. An alternate approach is provided by the Dimensions of Temperament Survery (Lerner, Palermo, Spiro, & Nesselroad, 1982). The scale was designed to be a measure of temperament from the preschool period through the adult level. The scale has three forms each consisting of the same 34 items phrased in different ways. The preschool form is a rating scale with items phrased in the following form: "My child falls asleep every night at the same time." The self-report scale for older children and the adult form phrase this item, "I fall asleep every night at the same time." A dichotomous response format is used on all three forms. The instrument was designed to test hypotheses about the effects of congruence and incongruence between parent and child temperament. For this reason an instrument producing the same factors at all ages was considered desirable. The research of the developers indicate that the five factors were found which were relatively invariant across the preschool to adult period; these were activity level, attention span, distractability, adaptability/approach, rhythmicity, and reactivity. Item selection, internal consistency and

factor structure based on a sizable sample are available, but the scale is in the formative stages so little validity or normative data is available.

Summary. Although a number of temperament measures have been developed, most were developed for specific research applications and there has been too little research using them to make definite statements about their validity. However, the Infant Temperament Questionnaire (Carey & McDevitt, 1978), and the Behavioral Style Questionnaire (McDevitt, 1978) have been extensively researched and have been shown to perform adequately. Unfortunately, this research has not been summarized and is widely scattered. The Temperament Assessment Battery (Martin, Note 3) was developed for clinical use. It is a three form battery, one for parents, one for teachers, and one for clinicians. The forms are accompanied by a manual which summerized reliability and validity data.

Educational and Clinical Correlates of Temperament

The following is a selected review of studies in four general areas: relationship of temperament and cognitive ability, achievement, psychopathology, and adjustment to early schooling experiences. These areas were chosen because they have educational and clinical significance and includes areas where there is active current research. Emphasis is given in this review to the most recent studies at the expense of much important earlier work. Further, the reviews are suggestive rather than exhaustive.

Relationship of Temperament and Cognitive Ability. Infant researchers have been actively investigating the relationship between temperament and cognitive development measured during the first year of life. For example, Lamb, Garn, and Keating (1981) obtained temperament and Bayley Mental Scale scores on over 33,000 eight month old children as part of the National Collaborative Perinatal Project. The temperamental variable of sociability was measured by three indices rated by psychologists after examining the infants. A low positive correlation was obtained between sociability and mental scale performance.

Vaughn, Taraldson, Crichton and Egland (1980) measured temperament using the Brazelton Neonatal device at seven and ten days of life and correlated five derived

temperamental like measures with Bayley Mental Scale scores obtained at nine months. Three of the five Brazelton measures correlated significantly and moderately with mental scores. In a similar study, Sostek and Anders (1977) obtained Brazelton measures at eight days, caretaker ratings of temperament on the Infant Temperament Scale (Carey, 1972) at 13 days, and Bayley scores at five to six months. Total Brazelton scores correlated .47 with mental scale performance, and two temperament variables (intensity and distractibility) correlated .58 and .55 with Bayley Mental scale performance.

There have also been several studies of the relationship between temperament and IQ measured during the preschool and early elementary school period. Martin (Note 2) found for 197 preschool children referred for a psychological evaluation that two parentally rated temperamental variables, persistence and adaptability, correlated significantly with IQ scores (.35 and .21, respectively).

Burk (1980) studied 125 children in nursery school through second grade who were attending a school for gifted children. All subjects had IQ scores in excess of 130 on the Wechsler Preschool and Primary Scale of Intelligence. Mother's of each child completed the Behavioral Style Questionnaire (McDevitt & Carey, 1978) When mean scores for this gifted group were compared to those of the norm sample of the Behavioral Style Questionnaire, it was found that the gifted group was rated significantly higher on approach/withdrawal, (high scores indicate a stronger approach tendency), adaptability, mood, and persistence. A significantly lower distractability score was obtained.

Gordon and Thomas (1967) sought to determine the relationship between temperament and teachers' appraisal of intelligence. The participants were 93 children in four kindergarten classes. Children were divided into four groups based on teacher ratings: (a) "plungers," children who jumped without hesitation into new situations; (b) "goalongers," children who did not plunge in, but positively adapted to new situations; (c) "sideliners," children who were removed from activities and only slowly joined in; and (d) "nonparticipators," children who remained removed from activities. These four groups were thought to be defined by the two temperamental characteristics of approach/withdrawal, and adaptability. At the end of the school year the teachers estimated the

intelligence of the children on a seven-point scale. The children's actual IQs were obtained several months later through a routine administration of the Kuhlman-Anderson Intelligence Test. It was determined that plungers were rated as more intelligent by the teachers than sideliners, despite equivalence on measured IQ. This result was interpreted as showing that the temperamental characteristics of adaptability and approach/withdrawal can negatively bias teachers' assessments of the capabilities of children who are low in adaptability and have a withdrawal tendency.

In a partial replication of the Gordon and Thomas study, Holbrook (1982) obtained temperamental ratings of 117 children from the first grade teachers on the adaptability and approach/withdrawal dimensions of the Parent Temperament Form (Martin, Note 3). Measures of IQ were obtained from the Otis-Lennon Ability Test, Elementary I Level. Several months after the temperament measures were obtained, teachers were asked to estimate the IQs of their students. A correlation of .50 was obtained between the adaptability rating and predicted IQ with variance due to measured IQ partialled out. A similar partial correlation ($r = .43$) was obtained for approach/withdrawal. This result strongly support the idea that teachers tend to overestimate the intellectual abilities of adaptable and socially approaching children and underestimate the intellectual abilities of less adaptable, withdrawing children.

This sample of studies demonstrates that tempermental variables, even when measured during the neonatal period, are related to subsequent measures of IQ or general cognitive ability. It is not clear at present what temperamental variables are the best predictors of cognitive ability, although variables related to sociability, including adaptability and approach/withdrawal, have been found to be related to cognitive ability in several studies. The mechanism by which sociability may be related to cognitive ability is hinted at by the results of the Gordon and Thomas (1967) and the Holbrook (1982) studies. That is, caretaker may form higher expectations for cognitive development for more sociable children than for less sociable children. Further, more sociable children may elicit more caretaker stimulation just because these children are more pleasant to be around. Further

research of this relationship between temperament and cognitive ability is obviously called for in order to clarify these mechanisms.

Relationship Between Academic Achievement and Temperament. Thomas and Chess and coworkers have carried out research on the relationship between temperament ratings and academic achievement for children in the New York Longitudinal Study. These correlations were between temperament ratings at age five and academic achievement scores in reading and arithmetic obtained at various times during the elementary school years. Achievement data consisted of all standardized tests administered in grades one through six. The Wide Range Achievement Test was also administered by the research staff to 79 of the children in the sample. Low adaptability and low approach/withdrawal ratings (indicating withdrawal were significant predictors of low achievement.

Burk (1980) studied the relationship between achievement and temperament for a group of gifted kindergarten, first and second grade children. Temperament was assessed by parent ratings on the Behavioral Style Questionnaire and achievement was measured by the reading and arithmetic subtests of the Stanford Achievement Test. Activity level and persistence produced a significant multiple R for reading achievement ($r = .29$). No temperament variable significantly predicted arithmetic achievement. The meaning of these findings are difficult to determine because this study suffered from methodological problems, including a restriction of range; i.e., all subjects were in gifted programs and were achieving at uniformily high levels.

Pullis and Cadwell (1982) studied the relationship between teachers' estimates of academic achievement and three temperamental factors (task orientation, reactivity, and adaptability) for a large sample of kindergarten, first, and second grade children. Academic achievement ratings were found to be significantly related to adaptability ($r = .44$) and task orientation ($r = .76$). Since the task orientation factor was constructed primarily from persistence and distractibility scales, this outcome supports the finding for persistence obtained by Burk. Adaptability and approach/withdrawal items had high loadings on the adaptability factor so this result supports the findings of Thomas and Chess (1977).

Another study supporting the importance of adaptability in academic achievement was carried out by Holbrook (1982). This investigator looked at the relationship between adaptability and approach/withdrawal, and achievement. Teacher ratings were obtained on the Teacher Temperament Form (Thomas & Chess, 1977--Martin Revision) for 117 first grade children as were teacher assigned grades. Also standardized achievement data from the American School Achievement Test, Primary Battery I (Revised Edition), and IQ estimates from the Otis-Lennon Ability Test, Primary I, were obtained. Adaptability was found to correlate significantly with both reading and math grades (.22 and .48 respectively) with variance due to IQ partialled out. Slightly lower but significant correlations were obtained between standardized achievement scores in reading and math (.35 and .38 respectively) with IQ partialled out. No significant relationship was found for the approach/withdrawal dimension.

Although every study did not find the same set of temperament variables related achievement, persistence and distractibility, and adaptability have been significantly related to achievement in several studies and seem to warrant the greatest attention in further studies of this relationship. The findings for sociability variables--adaptability and approach/withdrawal--leads to implication that children who are more pleasant in the view of the caretaker may be stimulated more by those caretakers and the positive social "halo" may create higher expectations for achievement. An alternative hypothesis that cannot be ruled out is that high achievers function at higher cognitive levels and that this ability is related to social intelligence or the ability to understand and appropriately relate to the social environment.

Relationship Between Temperament and Psychopathology. The most influential study of the relationship between temperament and psychotherapy was the New York Longitudinal Study of Thomas, Chess and colleagues (Thomas & Chess, 1977). These researchers obtained parental temperament measures on 141 children through parental interviews held on a three month basis during the first year of life and annually thereafter until the child was five years of age. Parents who participated in the research were informed that as a part of the study they could refer their children to the clinical director of the project if they

felt the child was manifesting symptoms of a behavioral or emotional problem. By the time the sample had reached five years of age 42 active clinical cases had been isolated. This is a cumulative number and not all were actively manifesting symptoms at any one time. When temperamental differences between the clinical and non-clinical groups were analyzed seven of the nine temperamental variables differentiated significantly between the groups. Further analysis demonstrated that both before and after the development of behavior problems the clinical groups differed from the non-clinical group in their temperament characterictics. A particularly strong associate was obtained for children manifesting the "difficult child" pattern of temperamental traits and clinical symptoms. The "difficult child" cluster of temperaments which were isolated by clinical as well as empirical means consisted of low biological rhythmicity, poor adaptability, low approach tendency, negative mood, and high emotional intensity. Seventy percent of the children with this temperamental pattern manifest some form of behavioral disturbance, while only 22 percent of the non-difficult children were found to have behavioral disturbances.

Graham, Rutter, and George (1973) studied 60 children who had at least one mentally ill parent. They obtained parental ratings of temperament, then one year later obtained ratings on a behavior questionnaire from parents and teachers. Using the responses to the behavior questionnaire to identify the maladjusted children, it was determined that high intensity, negative mood, low regularity and low fastidiousness (a temperament measure added to the Thomas and Chess list by these researchers) were related to maladjustment.

Terestman (1980) obtained nursery school teacher temperament ratings of a portion of the New York Longitudinal Study sample. She followed these children for five years documenting cases of clinically significant symptoms. She determined that mood and intensity as rated by nursery school teachers significantly discriminated between children manifesting symptoms and those not manifesting symptoms.

In a recent concurrent validity study Pfeffer and Martin (in press) obtained mother and father temperamental ratings on children referred for a psychological evaluation and a matched set of children who had not been referred. The referred children were rated as more active, less rhythmic, less adaptable, less persistent, and as having a higher stimulus threshold (less sensitive).

These studies demonstrate (a) that several temperament variables are related to the manifestation of maladaptive symptoms, the number and specific symptoms involved probably are related to the actual problems demonstrated by the children; (b) emotional intensity has been found to be associated with maladjustment in three of the four studies reviewed, and is probably the single best indicator of emotional difficulties; (c) three of the studies demonstrated that specific temperamental patterns predate the onset of symptoms.

Studies in the future must specify more precisely the behaviors demonstrated by the maladjusted groups so that temperament-behavior correlations can be more meaningfully determined. Lambert and Windmiller (1977) studied a carefully delineated group of hyperactive children and found that only distractibility differentiated this group from other groups. This study demonstrates a more precise connection between a specific temperamental variable and a specific behavior problem, and exemplifies the type of study that is needed in this area of inquiry.

There have been a number of attempts to determine the mechanisms or processes by which certain temperament patterns predispose children toward behavior problems. Baker (Note 5), for example, determined that two year olds who were rated by their parents as having the difficult child cluster of temperamental characteristics were significantly more fearful and exhibited more frustration than other children in several experimental situations. Bates (Note 6) noted that mothers perceive their difficult children age 6 to 13 months as demanding more attention and as being unable to entertain themselves. Unrelated mothers also perceive their cries to be more abrasive. Further, two year olds with difficult temperaments have more conflicted interactions with their mothers than easier toddlers and perceived difficultness from as early as age six months predicts maternal reports of behavior problems at age three years (Bates, Note 6).

This sketchy review provides the outlines of a mechanism for the development of maladjustment of difficult children. That is, the child's difficult behavior negatively effects parental behaviors and attitudes. These attitudes probably result in fewer or less skilled caretaker interactions during infancy and thereafter. Thus, the parental

behavior produces an environment that is riskier for the child because the behavior doesn't meet the child's needs in an adequate manner. These factors, then, raise the probability of maladjustment.

Relationship Between Temperament and Early School Adjustment. In the previous section we have seen that there is a good deal of evidence that temperamental variables are related to the development of psychopathological symptoms. While this research has important implications for child psychologists, this body of research does not address the equally important question of the relationship between temperament and adjustment within the "normal" range.

Four studies have been located which investigated the relationship between adjustment to school and temperament. Carey, Fox, and McDevitt (1977) obtained parental temperament ratings on the Behavioral Style Questionnaire (Carey & McDevitt, 1978) and teacher adjustment ratings on the Bommarito Socialization Scale for 51 first grade children. They found that only adaptability ratings were significantly correlated with school adjustment, and the relationship was modest ($r = .35$). The researchers suggested that this relatively weak relationship was due to inadequacies in the adjustment scale.

Burk (1980) studied 125 gifted children in kindergarten, first and second grades and obtained very similar results. Again parent temperament ratings were obtained on the Behavioral Style Questionnaire while school adjustment was rated by teachers on the Child's Behavior Trait Scale. A significant but low correlation was obtained between adaptability and adjustment ($r = .23$) and between persistence and adjustment ($r = .35$). The significance of persistence in this study was interpreted as relating to the importance placed on persistence in programs for the gifted.

Scholom, Zucker, and Stollack (1979) carried out a retrospective study of 132 three and four year olds. Parents were asked to rate the child's temperament as an infant and these ratings were compared to nursery school teacher ratings of adjustment. These researchers found that infant mood ratings (a factor made up of adaptability, approach/withdrawal, and mood) was the best predictor of adjustment.

In a study of adjustment to nursery school, Feuerstein and Martin (Note 7) obtained parental temperament ratings on the Martin revision of the Thomas and Chess Parent Temperament Questionnaire. These children were also rated by their nursery school teachers on the California Preschool Social Competence Scale (CPSCS), and on a one-item continuum designed to be a global measure of social adjustment to school. The children also participated in a sociometric procedure designed to measure popularity. Consistent with previous research adaptability was found to be significantly related to both teachers measures of adjustment ($r = .25$ with CPSCS, $r = .48$ with the one-item ratings, although not to the sociometric popularity measure. While other temperamental variables were found to relate to adjustment regression analysis demonstrated that these variables added little to the variance explained by adaptability.

Based on these studies it can be said that maternally rated temperament variables, primarily adaptability, account for from 10 percent to 25 percent of the variance in teacher rated adjustment to school in the early grades. These percentages at first glance seem disappointingly low. However, school adjustment is a very complex variable and temperamental variables could hardly be expected to be related to some subsets of behavior in this domain (e.g., verbal expression). It is known that measures in the cognitive domain (e.g., intelligence measures) correlate substantially even in kindergarten with such variables as verbal expression. Thus, the important issue which remains to be clarified in the extent to which temperamental variables add unique variance to the predictions of school adjustment over that variance predicted by other measures. If they do add unique variance, 10 to 25 percent of the variance explained could be very important and meaningful.

Conclusion

While temperamental variables have been the subject of speculation and some research in psychology for a long time, research based on contemporary theory is still in its infancy. This review has shown, however, that there are potentially important relationships between temperamental variables and such variables as cognitive ability, academic achievement, and school adjustment. Many of the studies reviewed have been longitudinal

investigations in which the place of temperamental variables in the development of cognitive ability and social-emotional maturity have been relatively clear. These are trends which child psychologists working in schools and clinics should be aware of and actively investigating.

Because the measures available for children are still in experimental stages of development, applied psychologists cannot simply incorporate one or more of them into their current assessment battery. However, the Infant Temperament Questionnaire (Carey & McDevitt, 1978), the Behavioral Style Questionnaire (McDevitt & Carey, 1978) and The Temperament Assessment Battery (Martin, Note 3) are of sufficient psychometric quality that they could be used in some screening programs as methods of systematically gathering information about the temperamental characteristics of children, if the limitations of the measure were kept in mind. The future for temperamental assessment and the general area of social-emotional developmental theory looks particularly bright. If current high levels of research activity continue it seems that very early and precise interventions into social-emotional developmental processes that have gone astray will become possible.

Reference Notes

1. Fullard, W., McDevitt, S. C., & Carey, W. B. Toddler Temperament Scale. Unpublished instrument, 1978.

2. Martin, R. P. A report on an item analysis, reliability, and validity study of the Thomas, Chess, and Korn Temperament Questionnaire Parent Form. Unpublished manuscript, 1979.

3. Martin, R. P. Manual for the Temperament Assessment Battery. Unpublished manuscript, 1982.

4. Hegvik, R. L., McDevitt, S. C., & Carey, W. B. Middle Childhood Temperament Questionnaire. Unpublished instrument, 1980.

5. Baker, E. J. Attachment and temperament correlates of two-year olds' reactions to strangers. Paper presented at the annual meeting of the American Psychological Association, Los Angeles, 1981.

6. Bates, J. E. Temperament as a part of social relationships: Implications of perceived infant difficultness. Unpublished paper.

7. Feuerstein, P., & Martin, R. P. Relationship between temperament and adjustment in nursery school. Manuscript submitted for publication, 1982.

References

Allport, G. W. Personality: A psychological interpretation. New York: Holt, 1937.

Als, H., Tronick, E., Lester, B. M., & Brazelton, T. B. Specific neonatal measure: The Brazelton Neonatal Behavior Assessment Scale. In J. D. Osofsky (Ed.), Handbook of of Infant Development. New York: Wiley, 1979.

Brazelton, T. B. Neonatal Behavioral Assessment Scale. Philadelphia: Lippincott, 1973.

Burk, E. Relationship of temperamental traits to achievement and adjustment in gifted children. Unpublished doctoral dissertation, Fordham University, 1980.

Buss, A. H., & Plomin R. A temperament theory of personality development. New York: Wiley, 1975.

Carey, W. B. A simplified method for measuring infant temperament. Journal of Pediatrics, 1970, 77, 188-194.

Carey, W. B. Measuring infant temperament. Journal of Pediatrics, 1972, 81, 414.

Carey, W. B., Fox, M., & McDevitt, S. C. Temperament as a factor in early school adjustment. Pediatrics, 1977, 60, 621-624.

Carey, W. B., & McDevitt, S. C. A revision of the Infant Temperament Questionnaire. Pediatrics, 1978, 61,

Cattell, R. B. Description and measurement of personality. New York: World Book, 1946.

Diamond, S. Personality and temperament. New York: Harper, 1957.

Eliasz, A. Temperament and transituational stability of behavior in the physical and and social environment. Polish Psychological Bulletin, 1980, 11, 1-22.

Goldsmith, H. H., & Campos, J. J. Toward a theory of infant temperament. In R. N. Ende and R. J. Harmon (Eds.), Attachment and affiliation systems: Neurobiological and aspects. New York: Plenum, 1981.

Goldsmith, H. H., & Gottesman, I. I. Origins of variations in behavioral style: A longitudinal study of temperament in young twins. Child Development, 1981, 52, 91-103.

Gordon, E. M., & Thomas, A. Children's behavioral style and the teachers' appraisal of their intelligence. Journal of School Psychology, 1967, 5, 292-300.

Graham, P., Rutter, M., & George, S. Temperamental characteristics as predictors of behavior disorders in children. American Journal of Orthopsychiatry, 1973, 43, 3, 328-339.

Holbrook, J. Pupil temperament characteristics, teacher appraisal of intelligence, and assigned grades. Unpublished doctoral dissertation, University of Georgia, 1982.

Hynd, G. W., & Obrzut, J. E. (Eds.). Neuropsychological assessment and the school-age child. New York: Grune & Stratton, 1981.

Lamb, M. E., Garn, S. M., & Keating, M. T. Correlations between sociability and cognitive performance among eight-month olds. Child Development, 1981, 52, 711-713.

Lambert, N. M., & Windmiller, M. An exploratory study of temperament traits in a population of children at risk. Journal of Special Education, 1977, 11, 37-47.

Lerner, R., Palermo, M., Spiro, A., & Nesselroade, J. R. Assessing the dimensions of temperamental individuality across the life span: The Dimensions of Temperament Survey (DOTS). Child Development, 1982, 53, 149-159.

Maccoby, E. E., Doering, C. H., Jacklin, C. M., & Kraemer, H. Concentrations of sex harmones in unibical-cord blood: Their relation to sex and birth order of infants. Child Development, 1979, 50, 632-642.

McDevitt, S. C., & Carey, W. B. The measurement of temperament in 3-7 year old children. Journal of Child Psychology and Psychiatry, 1978, 19, 245-253.

Pfeffer, J., & Martin, R. P. Comparison of mother's and father's temperament ratings of referred and non-referred preschool children. Journal of Clinical Psychology, (in press).

Pullio, M., & Cadwell, J. The influence of children's temperament characteristics on teachers decision strategies. American Educational Research Journal, 1982, 19, 2, 165-181.

Rothbart, M. K. Measurement of temperament in infancy. Child Development, 1981, 52, 569-578.

Rothbart, M. K., & Derryberry, O. Development of individual differences in temperament. In M. E. Lamb & A. L. Brown (Eds.), Advance in developmental psychology (Vol. 1). Hillsdale, N. J.: Lawrence Erlbaum, 1981.

Scarr, S. Social intraversion-extraversion as a heritable response. Child Development, 1969, 40, 823-832.

Scholom, A., Zucker, R. A., & Stollack, G. E. Relating early child adjustment to infant and parent temperament. Journal of Abnormal Child Psychology, 1979, 7, 297-308.

Sheldon, W. H. The varieties of temperament: A psychology of constitutional differences. New York: Harper, 1942.

Sostek, A. M., & Anders, T. F. Relationship among Brazelton neonatal scale, Bayley infant scales, and early temperament. Child Development, 1977, 48, 320-323.

Terestman, N. Mood quality and intensity in nursery school children as predictors of behavior disorder. American Journal of Orthopsychiatry, 1980, 50, 125-138.

Thomas, A., & Chess, S. Temperament and development. Brunner/Mazel: New York, 1977.

Thomas, A., & Chess, S. The dynamics of psychological development. New York: Brunner/Mazel, 1980.

Thomas, A., & Chess, S., Birch, H., Hertzig, M. E., & Korn, S. Behavioral individuality in early childhood. New York: New York University Press, 1963.

Thompson, J. L., & Fuller, W. R. Foundations of behavior genetics. St. Louis: Mosby, 1978.

Vaughn, B., Taraldson, B., Crichton, L., & Egland, B. Relationships between neonatal behavioral organization and infant behavior during the first year of life. Infant Behavior and Development, 1980, 3, 47-66.

Zuckerman, M. Sensation seeking: Beyond the optimal level of arousal. Hillsdale, N. J.: Lawrence Erlbaum, 1979.